**Crime Rate Data description**

This data set gives a variety of variables by US state at two time points 10 years apart. A variety of regressions and t-tests can be carried out with the main scale dependent being and Crime Rate (offences per million population) and t-tests with the independent being whether or not the state is in the south. Mostly discrete variables as they measure populations per 1000, there are some continuous variables such as those measuring expenditure.

|  |  |  |
| --- | --- | --- |
| Variable | Description |  |
| CrimeRate | Crime rate (number of offences per million population) | Continuous |
| Youth | Young males (number of males aged 18-24 per 1000) | Discrete |
| Southern | Southern state 1 = yes, 0 = no | Binary |
| Education | Education time (average number of years schooling up to 25) | Discrete |
| ExpenditureYear0 | Expenditure (per capita expenditure on police) **skewed** | Continuous |
| LabourForce | Youth labour force (males employed 18-24 per 1000) | Discrete |
| Males | Males (per 1000 females) | Discrete |
| MoreMales | More males identified per 1000 females 1 = yes, 0 = no | Binary |
| StateSize | State size (in hundred thousands) | Discrete |
| YouthUnemployment | Youth Unemployment (number of males aged 18-24 per 1000) **skewed** | Discrete |
| MatureUnemployment | Mature Unemployment (number of males aged 35-39 per 1000) | Discrete |
| HighYouthUnemploy | High Youth Unemployment 1 = yes, 0 = no (high if Youth >3\*Mature ) | Binary |
| Wage | Wage (median weekly wage) | Continuous |
| BelowWage | Below Wage (number of families below half wage per 1000) | Discrete |

Note: The same variables are collected 10 years later and have 10 on the end.

Can be used for :

|  |  |  |
| --- | --- | --- |
| Test | Example Questions |  |
| 1. Independent t-test | Do states in the south have a higher crime rate? |  |
| 1. Correlation | Relationship between crime rate and police expenditure (Spearman’s) | Relationship between crime rate and number of families below half wage (Pearson’s) |
| 1. Simple regression | Can the number of males to females predict the crime rate? | Can number of families below half wage predict crime rate? |
| 1. Multiple regression | Crime rate with any independent. |  |
| 1. Chi-squared | Is there a relationship between southern states and the male population? | Is there a relationship between high youth unemployment and southern states? |
| 1. Paired t-test | Have crime rates increased in 10 years? | Has adult unemployment dropped in 10 years? |
| 1. Checking normality | Is the crime rate normally distributed? | Is youth unemployment normally distributed? |
| 1. Mann-Whitney | Is police expenditure higher in the south? | Is youth unemployment higher in the south? |